Renewal stochastic processes with correlated events. Phase transitions along time evolution.

Jorge Velázquez[1] and Alberto Robledo[1,2,*]
[1]*Instituto de Física, Universidad Nacional Autónoma de México, Apartado postal 20-364, México 01000 D.F., Mexico*
[2]*Departamento de Matemáticas, Universidad Carlos III de Madrid, Spain* (*on sabbatical leave*)
(Dated: January 11, 2013)

We consider renewal stochastic processes generated by non-independent events from the perspective that their basic distribution and associated generating functions obey the statistical-mechanical structure of systems with interacting degrees of freedom. Based on this fact we look briefly into the less known case of processes that display phase transitions along time. When the density distribution $\psi_n(t)$ for the occurrence of the n-th event at time t is considered to be a partition function, of a 'microcanonical' type for n 'degrees of freedom' at fixed 'energy' t, one obtains a set of four partition functions of which that for the generating function variable z and Laplace transform variable ϵ, conjugate to n and t, respectively, plays a central role. These partition functions relate to each other in the customary way and in accordance to the precepts of large deviations theory, while the entropy, or Massieu potential, derived from $\psi_n(t)$ satisfies an Euler relation. We illustrate this scheme first for an ordinary renewal process of events generated by a simple exponential waiting time distribution $\psi(t)$. Then we examine a process modelled after the so-called Hamiltonian Mean Field (HMF) model that is representative of agents that perform a repeated task with an associated outcome, such as an opinion poll. When a sequence of (many) events takes place in a sufficiently short time the process exhibits clustering of the outcome, but for larger times the process resembles that of independent events. The two regimes are separated by a sharp transition, technically of the second order. Finally we point out the existence of a similar scheme for random walk processes.

PACS numbers: 02.50.-r, 05.20.-y, 05.70.Fh

arXiv:1102.0594v1 [cond-mat.stat-mech] 3 Feb 2011

I. INTRODUCTION

A large class of stochastic processes are renewal processes [1] [2]. This class of sequences are generally used to model independent identically distributed (iid) occurrences. The renewal processes are concerned with the times of substitution of components that are replaced as soon as they break down. Here we recall [3] [4] that this basic type of stochastic process possesses all the elements of a statistical-mechanical system and therefore can be couched into this language and benefit from well-established methods and applications developed for the study of systems with many degrees of freedom. The common iid process maps into the non-interacting case, but the most important potential application of this equivalence is to the generalization of renewal processes to correlated events, where the large body of knowledge accumulated in the study of (short or long range) interacting particle or spin systems can find interesting guidelines or clear-cut analogies for renewal processes. One particular property that we present here is the occurrence of phase transitions along time evolution.

The layout of the article is as follows. We start in Section 2 with a concise description of a renewal process that involves the transformation into Laplace space of the relevant probability density functions and the use of generating functions [5]. In Section 3 we make explicit the statistical-mechanical ensemble structure, with only two pairs of conjugate variables, of the renewal process and illustrate the form that the partition functions take for the simple case of an exponentially-decaying waiting-time distribution density. In Section 4 we apply the saddle-point approximation in the evaluation of the partition functions and show that the required Legendre transform structure, the associated equations of state, and the Euler relation [6] are present in the formalism for the renewal process. In Section 5 we consider a specific example of a renewal process with correlated events that exhibits a phase transition when the time variable increases. The renewal process is representative, for instance, of an opinion poll, and is constructed to be equivalent to the statistical-mechanical Hamiltonian Mean Field (HMF) model of interacting particles [7–9]. Finally in Section 6 we summarize and discuss our results.

II. BASICS OF RENEWAL PROCESSES

Technically, an *ordinary* renewal process is a sequence of partial sums of iid positive random variables. This process may be thought of as a sequence of points in time when the lifetimes of some objects of the same type ends and they are replaced by new ones. The renewal process counts the number of renewals in the interval $[0, t)$, hence such a renewal counting process is a random piecewise constant function. A convenient analytical procedure to determine the properties of this kind of process is that

* robledo@fisica.unam.mx

of Montroll [5]. It is resumed as follows: Let $\psi(t)$ be the (normalized) waiting time distribution density for a single event and $\psi_n(t)$ the distribution density for the occurrence of the n-th event at time t. For iid events these densities are linked via

$$\psi_n(t) = \int_0^t dt' \ \psi(t-t') \ \psi_{n-1}(t'), \ n > 1, \tag{1}$$

or in Laplace space by

$$\widehat{\psi}_n(\epsilon) = \left[\widehat{\psi}(\epsilon)\right]^n, \tag{2}$$

where

$$\widehat{\psi}(\epsilon) = \int_0^\infty dt \ \exp(-\epsilon t) \ \psi(t) \tag{3}$$

and

$$\widehat{\psi}_n(\epsilon) = \int_0^\infty dt \ \exp(-\epsilon t) \ \psi_n(t). \tag{4}$$

We shall consider throughout this paper time variables to be dimensionless. A generating function for the $\psi_n(t)$ is defined via the z-transform

$$\psi(t; z) \equiv \sum_{n=1}^{\infty} \psi_n(t) \ z^n, \tag{5}$$

so that

$$\widehat{\psi}(\epsilon; z) \equiv \int_0^\infty dt \ \exp(-\epsilon t) \ \psi(t; z) = \sum_{n=1}^{\infty} \widehat{\psi}_n(\epsilon) \ z^n. \tag{6}$$

Use of Eq. (2) above turns $\widehat{\psi}(\epsilon; z)$ into a geometric series that when convergent becomes

$$\widehat{\psi}(\epsilon; z) = \left[\widehat{\psi}(\epsilon) \ z\right] / \left[1 - \widehat{\psi}(\epsilon) \ z\right]. \tag{7}$$

The functions $\psi_n(t)$ and $\psi(t; z)$ are recovered from $\widehat{\psi}_n(\epsilon)$ and $\widehat{\psi}(\epsilon; z)$, respectively, via inverse Laplace and inverse z transforms. The average time between events, or period, T is given by the first moment of $\psi(t)$,

$$T \equiv \int_0^\infty dt \ t \ \psi(t) = -\frac{d}{d\epsilon} \ln \widehat{\psi}(\epsilon)\Big|_{\epsilon=0} < \infty, \tag{8}$$

whereas the average number of events $\langle n(t) \rangle$ of a renewal sequence when the last event occurs at time t [10] is

$$\langle n(t) \rangle \equiv \frac{\sum_{n=1}^{\infty} n \ \psi_n(t)}{\psi(t; 1)} = z \frac{d}{dz} \ln \psi(t; z)\Big|_{z=1}. \tag{9}$$

Therefore, if the most common calculation aim is to determine $\psi_n(t)$ or $\langle n(t) \rangle$ for any given waiting time $\psi(t)$ distribution, use of $\widehat{\psi}(\epsilon)$ in Eqs. (2) and (7) followed by inverse transformation is an expedient method.

Here we recapture [3] [4] a precise interpretation of the above expressions while calling attention that it is not restricted to iid processes. This is that the functions $\psi_n(t)$, $\widehat{\psi}_n(\epsilon)$, $\psi(t; z)$, and $\widehat{\psi}(\epsilon; z)$ can be seen to be partition functions associated to an equilibrium statistical-mechanical system of n degrees of freedom arranged in configurations with energy measured by a time t. Below we detail that in the large t and n limits these functions can be evaluated via the saddle-point approximation and that this central statistical-mechanical property leads to equations of state and entropies or free energies related via Legendre transforms, where the variables ϵ and $\mu \equiv \ln z$ appear to be conjugate to the variables t and n, respectively.

III. STATISTICAL ENSEMBLES FOR RENEWAL PROCESSES

We observe that Eq. (6), $\widehat{\psi}(\epsilon; z) = \sum_{n=1}^{\infty} \widehat{\psi}_n(\epsilon) \ z^n$, has the form of the expression for the grand canonical partition function of a thermal system if we were to consider that the number of events n represents the number of particles or degrees of freedom, ϵ the inverse temperature, z the activity, and therefore $\widehat{\psi}_n(\epsilon)$ plays the role of the canonical partition function. Having considered that $\widehat{\psi}_n(\epsilon)$, the Laplace transform of $\psi_n(t)$ in Eq. (4), plays this role, the formal analogy can be extended by identification of $\psi_n(t)$ as the microcanonical partition function where t is the energy. Further, the generating function $\psi(t; z)$ would then be seen as the partition function corresponding to an ensemble of fixed energy t and activity z. (For iid random variables $\widehat{\psi}_n(\epsilon)$ is given by Eq. (2) and the corresponding thermal system is made of identical non-interacting degrees of freedom with $\widehat{\psi}(\epsilon)$ the canonical partition function per degree of freedom).

The scope of this analogy can be further assessed by defining the following entropies or Massieu potentials [6],

$$\begin{aligned} S_{\epsilon,\mu} &\equiv \ln \widehat{\psi}(\epsilon; z), \ \ S_{\epsilon,n} \equiv \ln \widehat{\psi}_n(\epsilon), \\ S_{t,\mu} &\equiv \ln \psi(t; z), \ \ S_{t,n} \equiv \ln \psi_n(t), \end{aligned} \tag{10}$$

where we have introduced the 'chemical potential' $\mu \equiv \ln z$. (These quantities may be negative since the arguments of the logarithms may be less than unity. Notice that these arguments are probability densities or their Laplace and/or z-transforms, while in ordinary statistical mechanics the arguments are configuration numbers or their transforms). If for large n a thermodynamic limit or a large deviations property [11] arises, then these potential functions would be related via Legendre transforms involving the pairs of conjugate variables (n, μ) and (t, ϵ) and mediated via the corresponding equations

of state. An Euler relation of the type

$$S_{t,n} = t\epsilon - n\mu \tag{11}$$

would hold, and attention should be paid in the evaluation of $S_{\epsilon,\mu} = \ln \widehat{\psi}(\epsilon; z)$ as a cursory inspection of repeated Legendre transforms would imply $S_{\epsilon,\mu} = S_{t,n} - t\epsilon + n\mu = 0$. Below we show that Eq. (11) holds with a nonzero $S_{\epsilon,\mu}$.

To help us examine the validity of this formal structure in the following section we determine the above partition functions for the particular iid random variable case of an exponential waiting-time density $\psi(t) = b\exp(-bt)$. One obtains

$$\widehat{\psi}(\epsilon; z) = bz\,(b + \epsilon - bz)^{-1}\,, \tag{12}$$

$$\widehat{\psi}_n(\epsilon) = b^n\,(b + \epsilon)^{-n}\,, \tag{13}$$

$$\psi(t; z) = bz\exp(-bt + btz) \tag{14}$$

and

$$\psi_n(t) = \frac{(bt)^n}{n!}\exp(-bt), \tag{15}$$

where we recognize in the last equation the distribution density of a Poisson process.

IV. ANALOGY WITH STATISTICAL MECHANICS

The asymptotic solution of $\widehat{\psi}_n(\epsilon)$ for $n >> 1$ can be found by use of the steepest-descent approximation of the inverse z-transform of $\widehat{\psi}(\epsilon; z)$,

$$\widehat{\psi}_n(\epsilon) = \frac{1}{2\pi i}\oint dz\,\exp\left[(n-1)\left(-\ln z + (n-1)^{-1}\ln\widehat{\psi}(\epsilon; z)\right)\right], \tag{16}$$

One obtains

$$\ln\widehat{\psi}_n(\epsilon) \simeq -(n-1)\ln z_0 + \ln\widehat{\psi}(\epsilon; z_0), \tag{17}$$

where z_0 can be eliminated in favor of n via the steepest-descent condition

$$n - 1 = z\frac{d}{dz}\ln\widehat{\psi}(\epsilon; z)\bigg|_{z=z_0}. \tag{18}$$

It is then possible to write

$$\widehat{\psi}_n(\epsilon) = \exp S_{\epsilon,n}, \tag{19}$$

where $S_{\epsilon,n} = \ln\widehat{\psi}_n(\epsilon)$ is the Legendre transform $S_{\epsilon,n} = -(n-1)\mu_0 + S_{\epsilon,\mu_0}$ of $S_{\epsilon,\mu_0} = \ln\widehat{\psi}(\epsilon; z_0)$. For the exponential waiting-time density $\psi(t) = b\exp(-bt)$ this transformation leads to the equation of state

$$n - 1 = 1 + bz_0(b + \epsilon - bz_0)^{-1}, \tag{20}$$

and to the Massieu potential

$$S_{\epsilon,n} = \ln\left[\frac{b^{n-1}}{(b+\epsilon)^{n-1}}\frac{(n-1)^{n-1}}{(n-2)^{n-2}}\right]. \tag{21}$$

In the limit $n \to \infty$ Eq. (13) is recovered.

Similarly, the asymptotic solution of $\psi_n(t)$ for $n >> 1$ is obtained with the use of the steepest-descent approximation of the inverse Laplace transform of $\widehat{\psi}_n(\epsilon)$,

$$\psi_n(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} d\epsilon\,\exp\left[n\left(\epsilon\tau + \ln\widehat{\psi}(\epsilon)\right)\right], \tag{22}$$

where $\tau \equiv t/n$. One obtains

$$n^{-1}\ln\psi_n(n\tau) \simeq \epsilon_0\tau + \ln\widehat{\psi}(\epsilon_0),\ \ n >> 1, \tag{23}$$

where ϵ_0 can be eliminated in favor of τ via the steepest-descent condition

$$\tau = -\frac{d}{d\epsilon}\ln\widehat{\psi}(\epsilon)\bigg|_{\epsilon=\epsilon_0}. \tag{24}$$

In taking the limit $n \to \infty$ also $t \to \infty$ but τ is kept finite. We can therefore write

$$\psi_n(t) = \exp S_{t,n}, \tag{25}$$

where $S_{t,n} = \ln\psi_n(t)$ is the Legendre transform $S_{t,n} = -t\epsilon_0 + S_{\epsilon_0,n}$ of $S_{\epsilon_0,n} = \ln\widehat{\psi}_n(\epsilon_0)$. For exponential waiting times $\psi(t) = b\exp(-bt)$ this transformation leads to the equation of state

$$t = n(b + \epsilon_0)^{-1}, \tag{26}$$

and to the Massieu potential

$$S_{t,n} = \ln\left[\left(btn^{-1}\right)^n\exp(n)\exp(-bt)\right]. \tag{27}$$

Therefore Eq. (15) is recovered in the limit $n \to \infty$ (when we notice that the Stirling approximation of the factorial is part of Eq. (27)).

Lastly, following an analogous procedure the asymptotic form for $\psi(t; z)$ for $n >> 1$ is given by

$$\psi(t; z) = \exp S_{t,\mu}, \tag{28}$$

where the Massieu potential $S_{t,\mu}$ for $\psi(t) = b\exp(-bt)$,

$$S_{t,\mu} = \ln\left[btz\exp\left(1 + bt + btz\right)\right], \tag{29}$$

is obtained as the Legendre transform $S_{t,\mu} = t\epsilon_0 + S_{\epsilon_0,z}$ of $S_{\epsilon_0,\mu} = \ln\widehat{\psi}(\epsilon_0; z)$ with

$$t = (b + \epsilon_0 - bz)^{-1}. \tag{30}$$

Eq. (29) is consistent with Eq. (14) in the limit $n \to \infty$.

To make explicit the observance of the Euler relation Eq. (11) we note that the inverse Legendre transform

that yields $S_{\epsilon,\mu}$ from S_{ϵ,n_0}, $S_{\epsilon,\mu} = (n_0 - 1)\mu + S_{\epsilon,n_0}$, requires

$$\mu = -\frac{d}{d(n-1)} n \ln \widehat{\psi}(\epsilon)\Big|_{n=n_0} = -\ln \widehat{\psi}(\epsilon), \tag{31}$$

so that $S_{\epsilon,n_0} = n_0 \ln \widehat{\psi}(\epsilon) = -n_0\mu$ and $S_{\epsilon,\mu} = (n_0 - 1)\mu + S_{\epsilon,n_0} = -\mu$. This leads to

$$S_{t_0,n_0} = t_0\epsilon - (n_0 - 1)\mu + S_{\epsilon,\mu} = t_0\epsilon - n_0\mu. \tag{32}$$

We note that the existence of the Euler relation for a system with only two pairs of conjugate variables does not imply the vanishing of the thermodynamic potential, $S_{\epsilon,\mu}$, associated to two consecutive Legendre transforms of the basic potential, $S_{t,n}$, a homogeneous function of order one in both variables t and n [6]. Notably, the partition function $\widehat{\psi}(\epsilon; z)$ associated to the variables ϵ and μ remains a fundamental and most useful quantity for the description of the renewal process.

V. AN EXAMPLE OF A RENEWAL PROCESS WITH CORRELATED EVENTS

As an illustration of how developments in the statistical mechanics of interacting particle or spin systems may have meaningful translations to renewal processes we present here features of a renewal process with correlated events. We take inspiration from the so-called Hamiltonian Mean Field (HMF) model [7]-[9] to point out the occurrence of phase transitions along time evolution.

Consider a sequence of events, each of which, besides taking place at a given time t, assigns values to two scalar quantities τ and θ, the first within the time interval $0 \leq \tau \leq t$ taken by the event, and the second within a fixed finite interval, say $0 \leq \theta \leq 2\pi$. For instance, the process may represent an agent (or agents) that performs a repeated task with outcome (τ, θ) that is not independent of those for all the previous events. A string of such n events is described by the sequence of triplets $[(t_1; \tau_1, \theta_1), ..., (t_n; \tau_n, \theta_n)]$. The two collections of values $(\tau_1,\ \tau_2,...,\ \tau_n)$ and $(\theta_1,\ \theta_2,...,\ \theta_n)$ are used to construct two additional time variables, T_n, the "idle" time, and W_n, the "active" time, respectively, that together comprise the total time taken by the sequence of n events, i.e.

$$t_n = T_n + W_n. \tag{33}$$

The idle time T_n is simply given by

$$T_n = \sum_{i=1}^{n} \tau_i, \tag{34}$$

whereas the active time W_n measures the dispersion of the values $(\theta_1,\ \theta_2,...,\ \theta_n)$ over the unit circle, being large when these are spread out over $(0, 2\pi)$ and small when they concentrate around a given θ. Although there are many options to define W_n, for definiteness we chose it to be determined by

$$W_n = \frac{1}{2n} \sum_{i,j=1}^{n} \left[1 - \cos(|\theta_i - \theta_j|)\right] \leq t, \tag{35}$$

where all pairs $(\theta_i,\ \theta_j)$ are equally considered. (We recall that time variables are considered dimensionless). Clearly, the condition Eq. (33) imposes a restriction in the possible values of the sequences $(\tau_1,\ \tau_2,...,\ \tau_n)$ and $(\theta_1,\ \theta_2,...,\ \theta_n)$. As a more specific illustration of this kind of process let us suppose there is an opinion poll organization that sends an agent (or group of pollsters) to take a survey that consists of a succession of n completed questionnaires obtained in the time interval $(0, t)$. Each respondent has a tag θ that quantifies a characteristic of the population surveyed, such as age, race, home environment, etc., and therefore W_n reflects the degree of coverage bias, in, for instance, the consideration of young voters, minorities, or rural areas. The time τ associated to each respondent measures wasted time in collecting opinions, since some people do not answer calls, or refuse to answer the poll, or do not give candid answers, and consequently T_n represents the extent of nonresponse bias.

The probability density of occurrence of the n-th event at time t with outcome (θ, τ), $\psi_n(t; \tau, \theta)$, is evaluated in terms of the statistics of occurrence of the previous $n - 1$ events. This is best prescribed in terms of the Laplace transform of $\psi_n(t; \tau, \theta)$ with respect to t, $\widehat{\psi}_n(\epsilon; \tau, \theta)$. Specifically, the renewal process is given by

$$\widehat{\psi}_n(\epsilon; \tau, \theta) \equiv \exp\left[-\epsilon \left\langle \delta t_n(\tau, \theta)\right\rangle_{n-1}\right], \tag{36}$$

where the average $\langle ... \rangle_{n-1}$ is performed over all possible values of $(\tau_1, \theta_1), ..., (\tau_{n-1}, \theta_{n-1})$, and

$$\delta t_n(\tau, \theta) = T_n + W_n - T_{n-1} - W_{n-1}. \tag{37}$$

The analogy with the HMF model becomes evident when it is seen that Eq. (36) corresponds to Widom's particle insertion formula when applied to the thermal system (for vanishing chemical potential μ) [12] [13]. The roles of the number of particles, their positions (in the unit circle), inverse temperature, energy, kinetic energy, and potential energy of the HMF model, are given, respectively, by n, θ_i, ϵ, t_n, T_n, and W_n. (With no loss of generality a coupling constant in the potential energy term of the ferromagnetic HMF model has been set equal to unity). In our notation, the Helmholtz free energy of the HMF model in the limit $n \to \infty$, obtained via the saddle-point approximation [7]-[9], is

$$-\epsilon\widehat{\psi}_n(\epsilon) = -\frac{1}{2}\ln\left(\frac{\epsilon}{2\pi}\right) - \frac{\epsilon}{2} + \max_x\left(-\frac{\epsilon x^2}{2} + \ln 2\pi I_0(\epsilon x)\right) \tag{38}$$

where the auxiliary variable x satisfies

$$x = \epsilon \frac{I_1(\epsilon x)}{I_0(\epsilon x)} \tag{39}$$

and where $I_i(y)$ is the modified Bessel function of order i. As known [7]-[9], the HMF model exhibits two equilibrium phases, indicated by the possible solutions of Eq. (39). When $\epsilon < \epsilon_c = 2$ the variable x, identified as the model's magnetization M, vanishes, but x is nonzero for $\epsilon > \epsilon_c$, increasing gradually as ϵ increases and reaching unity as $\epsilon \to \infty$. These properties imply that the mean active time $\langle W_n \rangle$, the average of W_n over all sequences $(\tau_1, \theta_1), ..., (\tau_n, \theta_n)$, is given by $\langle W_n \rangle = (1 - x^2)n/2$, and that the relationship between the time t and the amplitude ϵ (the caloric equation for the HMF model) reads

$$t = \frac{n}{2\epsilon} + \langle W_n \rangle \,, \tag{40}$$

[7]-[9]. Thus, $\langle W_n \rangle$ displays a fixed maximum value

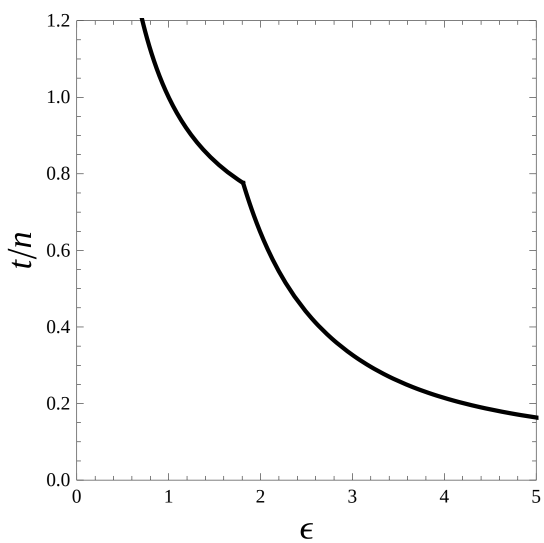

FIG. 1. Dependence of t/n, $t, n \gg 1$, on the Laplace variable ϵ for the renewal process model of correlated events designed to be analogous to the HMF model. The figure is equivalent to the caloric equation of the HMF model and shows the two-phase behavior described in the text.

$\langle W_n \rangle = n/2$ for $\epsilon < \epsilon_c = 2$, whereas it decreases and approaches zero as $\epsilon \to \infty$. The two-phase behavior and its transition at ϵ_c is reflected by the θ-dependence of $\widehat{\psi}_n(\epsilon; \tau, \theta)$ when $n \gg 1$. For small ϵ the distribution is uniform in θ, but when $\epsilon > \epsilon_c$ it becomes peaked around a given (although arbitrary) value of $\theta = \phi$. As we see below this feature is preserved in the distribution for the original variable t, i.e. there is a critical time t_c above which $\psi_n(t; \tau, \theta)$ is uniform in θ and below which it is peaked around a given $\theta = \phi$. The τ-dependence of $\widehat{\psi}_n(\epsilon; \tau, \theta)$ has an exponential form (Gaussian if written for the 'momentum' $p = \pm\sqrt{2\tau}$) for all ϵ. When $n \to \infty$ Eq. (36) leads to [9]

$$\widehat{\psi}_n(\epsilon; \tau, \theta) = \sqrt{\frac{\epsilon}{2\pi}} \exp\left(-\epsilon\tau\right) \frac{1}{2\pi I_0(\epsilon x)} \exp(\epsilon \mathbf{x} \cdot \theta), \tag{41}$$

where $\mathbf{x} = (x\cos\phi, x\sin\phi)$ and $\theta = (\cos\theta, \sin\theta)$.

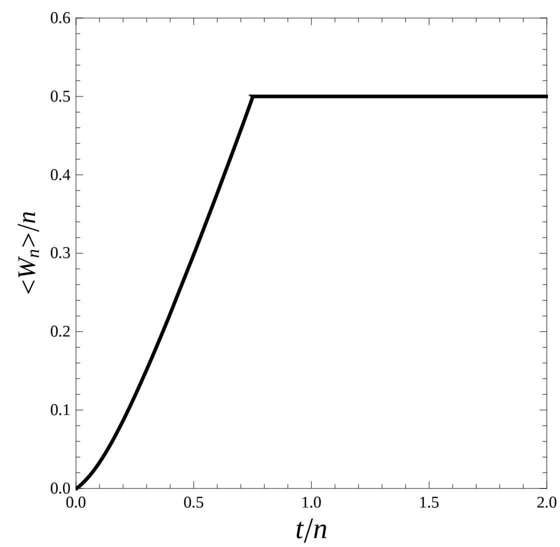

FIG. 2. Dependence of the active time per unit event $\langle W_n \rangle / n$ for $n \gg 1$ on t/n for the renewal process model of correlated events designed to be analogous to the HMF model. The figure shows the two-phase behavior described in the text.

The corresponding expression for $\psi_n(\epsilon; \tau, \theta)$, the inverse Laplace transform of Eq. (41) obtained via the saddle-point approximation, is

$$\psi(t; \tau, \theta) \simeq C \exp\left(x^2 - 1/2\right) \left(\frac{x^2 - 1/2}{t - \tau + \mathbf{x} \cdot \theta}\right)^{1/2} \left(I_0\left[\frac{(x^2 - 1/2)x}{t - \tau + \mathbf{x} \cdot \theta}\right]\right)^{-1}, \tag{42}$$

where C is a normalization constant. Following Refs. [8] [9] we evaluated the dependence of $\langle W_n \rangle / n$ on ϵ after solving numerically Eq. (39). Subsequently we used this in Eq. (40) to obtain the dependence of t/n on ϵ (shown in Fig. 1, where the two-phase feature is evident). This allowed us to determine the time dependence of the mean active time per event, $\langle W_n \rangle / n$, shown in Fig. 2, where it is observed that this quantity increases monotonically with t until it saturates at a value of $1/2$ at t_c and remains constant thereafter. Since $\langle W_n \rangle / n$ measures the average spread of the tags θ_i, $i = 1, ..., n$, we conclude that for short times $t < t_c$ this spread falls below its

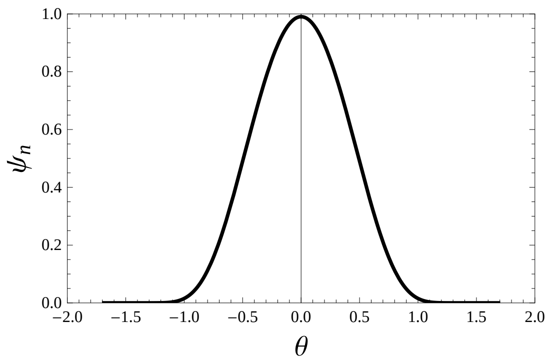

FIG. 3. Dependence of $\psi_n(t;\tau,\theta)$, $n \gg 1$, on the θ when $t < t_c$. When $t > t_c$ the function $\psi_n(t;\tau,\theta)$ is θ-independent.

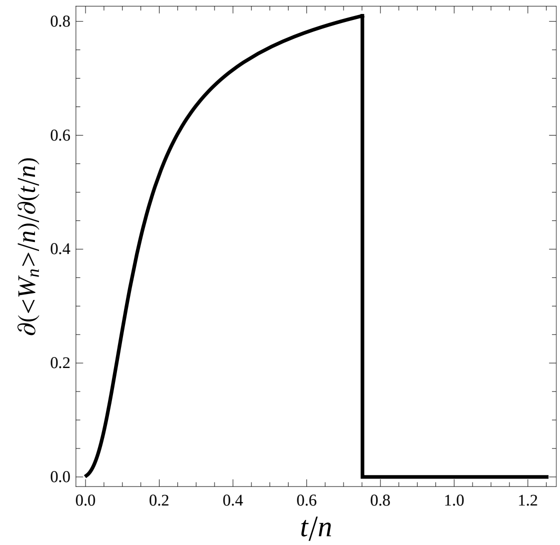

FIG. 4. Dependence of $\partial(\langle W_n\rangle/n)/\partial(t/n)$, $t, n \gg 1$, on t/n.

maximum whereas for larger times $t > t_c$ the maximum spread is always assured. In terms of the opinion poll sets of n samples taken inside time intervals $t < t_c$ suffer from coverage bias but are as free of it as it is possible when the set of samples are collected within time intervals $t > t_c$. This feature is corroborated in Fig. 3 where we show the θ-dependence of $\psi_n(t;\tau,\theta)$, $t < t_c$, as given by Ec.(42). When $t > t_c$ the density is flat, independent of θ. Interestingly, sequences of n events that take place within time intervals $t < t_c$ are correlated while those for $t > t_c$ are not. When $\langle W_n\rangle/n < 1/2$ there is on average no sufficient time for the pollster to move to other locations or to switch to different population groups, there is a repetition, or ordering in the set of samples. This generates a coverage bias. Similar arguments can be elaborated in terms of the average idle time $\langle T_n\rangle/n = 1/(2\epsilon)$ that in the example of an opinion poll is reflected by the presence of nonresponse bias. A measure of the correlations induced for $t < t_c$ is given by the time derivative of $\langle W_n\rangle/n$ (one of several response functions) as shown in Fig. 4. For $t > t_c$ the HMF model behaves effectively as an ideal gas, and, as we can see from Figs. 2 to 4, the renewal process conforms to that of independent events for this regime.

Thus, by construction our opinion poll renewal process acquires all the properties of the HMF model, mainly its second order phase transition that separates two different regimes. That is, for small t/n strings of n events cluster around a given value of the tag θ symptomatic of an inefficient poll, but for larger t/n the events display a uniform dispersal of θ suggesting the proper working of the sampling process. The clustering of the tag θ when $\epsilon > \epsilon_c$ is expressed by the Laplace transform variable ϵ as it measures the width of $\widehat{\psi}_n(\epsilon;\tau,\theta)$. There are other known interesting properties of the HMF model such as the occurrence of long lived, or quasistationary, states for temperatures below the transition temperature, when the system displays features of the high-temperature phase uniform in θ [8] [9]. These states would manifest also in the renewal process as sequences of active times $\langle W_n\rangle/n$ larger than those shown in Fig. 2 for some range of values $t < t_c$ close to t_c.

VI. SUMMARY AND DISCUSSION

We have made use of the statistical-mechanical interpretation of the basic elements that constitute the theory of renewal processes. Our purpose for recapturing this analogy is to facilitate the application of useful techniques and approximations built up and tested through a large amount of studies of thermal systems. Potentially these methodologies can have important effects in the study of complex systems that originate outside ordinary statistical-mechanical physical systems, in a variety of fields, in ecology, economy, sociology, etc. where stochastic processes such as that for the renewing of events often arise. The known parallels between renewal processes and statistical mechanics are an indication of the general, Laplace and Legendre transform structure of large deviations theory [11]. The saddle-point approximation is central to this theory where a probability P_n obeys the form $P_n \simeq \exp(-ns)$ for $n \gg 1$ with s a positive quantity independent of n named the rate function [11]. Clearly, the Massieu potentials $S_{\epsilon,n}$ and $S_{t,n}$ in Eqs. (19) and (25), respectively, when written as $S_{\epsilon,n} = -ns(\epsilon)$ and $S_{t,n} = -ns(\tau)$ comply with this property. Thus we could describe renewal processes, familiar in the probability theory domain, in this alternative language. Nevertheless, because of our stated purposes we have used a statistical-mechanical language. We have pointed out the equivalence with large deviation theory when appropriate. As a difference from the present statistical study of single sequences of correlated events, sets of renewal sequences dependent on each other (when uncoupled the sequences are made of uncorrelated events) have been analyzed with the use of multivariate distributions. For a recent application see Ref. [14].

To exemplify the use of the parallelism between renewal processes and statistical mechanics we devised a model renewal process with correlated events that displays a phase transition as time progresses. When n events take place within a relatively short time interval their correlation is evident, whereas for longer time intervals their statistical properties are identical to those of independent events, and there is a sharp transition between the two regimes. The renewal process ensemble structure facilitated the description of this model that we chose to portray, amongst several possible options, in terms of a polling process, and with characteristics taken straightforwardly from a well-known particle or spin statistical-mechanical model, the HMF model [7]. There are examples of phase transitions occurring along time evolution in deterministic (as opposed to stochastic) systems. See Ref. [15] and references therein.

We close by mentioning that correlated random walk processes on regular lattices, as described with the help of the Fourier transform and generating functions [5], and for both discrete and continuous time distributions, can be couched into a partition function language just as we have shown here for correlated renewal processes. Due to the sign of the integers used to locate the walker in lattice space the ensuing statistical-mechanical formalism differs also from the canonical type, basically by using the velocity instead of the kinetic energy as the primary variable that describes interacting particles.

ACKNOWLEDGMENTS

AR acknowledges an interesting conversation with F. Baldovin. We are grateful for support from DGAPA-UNAM and CONACyT (Mexican agencies). AR acknowledges support from MEC (Spain).

[1] S.I. Resnick, *Adventures in Stochastic Processes* (Birkhaüser, Boston, 1992).
[2] R.G. Callager, *Discrete Stochastic Processes* (Kluwer, MA, USA, 1996).
[3] R. Fowler, *Statistical Mechanics* (2nd edition) (Cambridge University Press, 1966).
[4] N.G. van Kampen, *Stochastic Processes in Physics and Chemistry* (2nd edition) (North-Holland, Amsterdam, 1991).
[5] E.W. Montroll, G.H. Weiss, J. Math. Phys., **6**, 167 (1965).
[6] H.B. Callen, *Thermodynamics and an Introduction to Themostatistics* (2nd edition) (John Wiley & Sons, New York, 1985).
[7] A. Campa, Th. Dauxois, S. Ruffo, Phys. Rep., **480**, 57 (2009).
[8] M. Antoni, S. Ruffo, Phys. Rev. E, **52**, 2361 (1995).
[9] V. Latora, A. Rapisarda, S. Ruffo, Physica D, **131**, 38 (1999).
[10] The density $\chi_n(t)$ for the occurrence of n events up to time t is obtained form $\psi_n(t)$ by consideration of the probability $\Psi(t) = 1 - \int_0^t dt'\ \psi(t')\ = \int_t^\infty dt'\ \psi(t')$ that there is no renewal event in the time interval $(0,t)$. In Laplace space one obtains $\widehat{\chi}_n(\epsilon) = \epsilon^{-1}(1 - \widehat{\psi}(\epsilon))\widehat{\psi}_n(\epsilon)$.
[11] H. Touchette, Phys. Rep., **478**, 1 (2009).
[12] B. Widom, J. Stat. Phys., **19**, 563 (1978).
[13] F. Baldovin, A. Robledo, in preparation.
[14] U. Sumita, J. Zuo, Oper. Res. Soc. Jpn., **53**, 119 (2010).
[15] C. Beck, F. Schlög, *Thermodynamics of Chaotic Systems* (Cambridge University Press, Cambridge, 1993) pp. 248–251.